

VU Research Portal

Effectonderzoek in de alternatieve gezondheidszorg: randomiseren en blinderen blijven integrale onderdelen

Biemans, M.C.J.; Bouter, L.M.; Knipschild, R.

published in

Nederlands Tijdschrift voor Integrale Geneeskunde
1987

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Biemans, M. C. J., Bouter, L. M., & Knipschild, R. (1987). Effectonderzoek in de alternatieve gezondheidszorg: randomiseren en blinderen blijven integrale onderdelen. *Nederlands Tijdschrift voor Integrale Geneeskunde*, 4, 159-168.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl



EFFECTONDERZOEK IN DE ALTERNATIEVE GEZONDHEIDSZORG

RANDOMISEREN EN BLINDEREN BLIJVEN INTEGRALE ONDERDELEN

DRS. M.J.C. BIEMANS, DRS. L.M. BOUTER
EN PROF.DR.P.G. KNIPSCHILD

INLEIDING

Al geruime tijd wordt er in Nederland op verschillende niveau's gediscussieerd over de vraag op welke wijze alternatieve geneeswijzen op hun effectiviteit onderzocht dienen te worden. Daarbij kan men constateren, dat veel onderzoeksvoorstellen die op een reguliere denkwijze berusten, door het alternatieve veld worden afgewezen. Recentelijk is dit ook naar voren gekomen in een tweetal artikelen in dit tijdschrift (1,2). In beide artikelen wordt een pleidooi gehouden voor een niet-gerandomiseerde onderzoeksopzet. Soms neemt men in alternatieve kringen zelfs genoegen met een ongecontroleerde onderzoeksopzet (3).

Hiertegenover staat dat in het reguliere circuit soms al te snel over placebo-effecten en zelfs kwakzalverij wordt gesproken door artsen die geen weet hebben van de bestaande wetenschappelijke literatuur, aangaande het effect van alternatieve geneeswijzen. Een analyse van publicaties die inmiddels op dit terrein zijn verschenen, laat zien dat het wel degelijk mogelijk is binnen de alternatieve geneeswijzen gecontroleerde experimenten uit te voeren.

Het in dit artikel ingenomen standpunt luidt, dat effectenonderzoek bij voorkeur experimenteel dient te geschieden. Dit geldt, naar de mening van de auteurs, in gelijke mate voor alternatieve en voor reguliere behandelingen. Het betoog is als volgt opgebouwd.

Allereerst wordt nagegaan, uit welke componenten het effect van een behandeling bestaat. Vervolgens

wordt een drietal bij het opzetten van effectonderzoek cruciale vragen geformuleerd en uitgewerkt. Tot slot wordt stil gestaan bij de noodzaak van randomiseren en blinderen. De theoretische uiteenzetting wordt geïllustreerd met voorbeelden van onderzoek naar de effectiviteit van alternatieve behandelwijzen die (ten dele) voldoen aan deze onderzoekscriteria.

EFFECTONDERZOEK

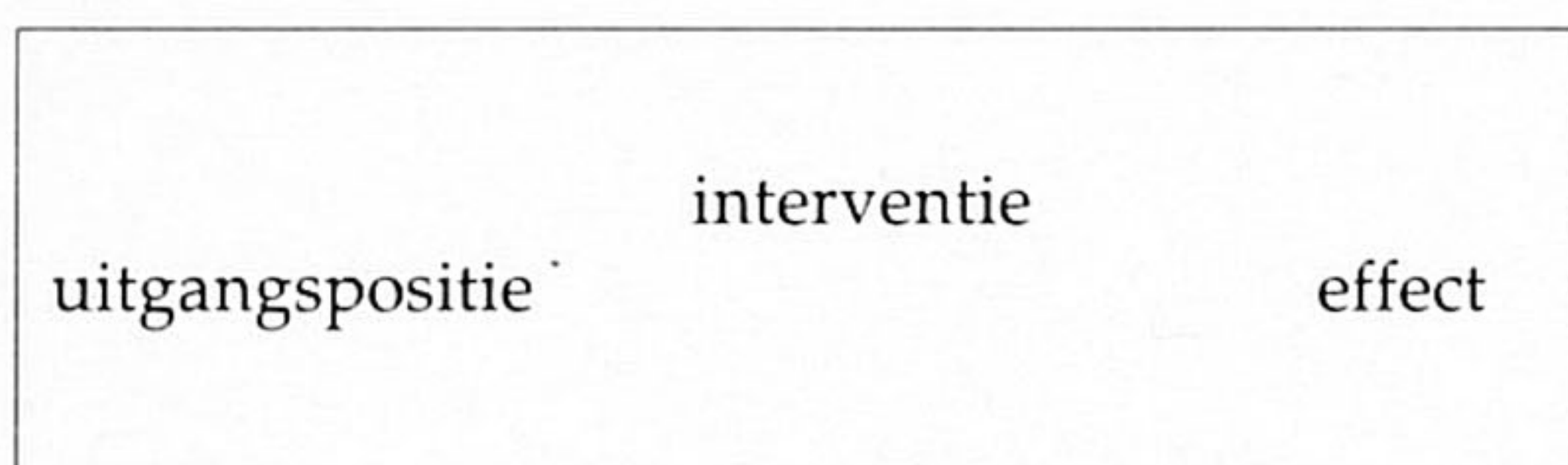
Zowel in de reguliere als de alternatieve gezondheidszorg voeren behandelaars interventies uit bij patiënten. Dit gebeurt in de regel, in de verwachting dat deze een gunstige wending geven aan het beloop van een aandoening en de daarmee samenhangende klachten. De aard van de interventies kan verschillend zijn. Als voorbeeld gebruiken we hier een patiënt met chronische lage rugpijn. Een behandelaar kan deze patiënt aanraden uit te kijken naar ander werk dat minder belastend is voor zijn rug. Er kan een recept uitgeschreven worden voor pijnstillende middelen of voor bepaalde oefeningen. De interventie zou ook kunnen bestaan uit manipulatie door een manueel therapeut, massage

Drs. M.J.C. Biemans, Drs. L.M. Bouter en Prof.dr.P.G. Knipschild zijn respectievelijk wetenschappelijk medewerkster, universitair docent en hoogleraar bij de Vakgroep Epidemiologie van de Rijksuniversiteit Limburg.



door een fysiotherapeut, stimulatie door een acupuncturist of een operatie uitgevoerd door een orthopedisch chirurg.

Het verband tussen interventie en effect vormt de centrale vraag voor zowel behandelaars als onderzoekers. Men dient zich daarbij niet tevreden te stellen met een constatering dat een bepaalde interventie in de dagelijkse praktijk doorgaans een bepaald effect oplevert. Men wil graag meer zekerheid hebben. Daarbij kan men zich bovendien afvragen, welke behandeling voor dergelijke patiënten het meest gunstige resultaat geeft. Door middel van een interventie te relateren aan veranderingen in een of meer effectparameters. De centrale vraag daarbij is niet hoe een eventueel effect tot stand komt, maar of, en in welke mate een interventie effect heeft. In figuur 1 is dit als volgt weer te geven.



figuur 1.

Bij het formuleren van een probleemstelling, waarop men door middel van een effectonderzoek antwoord tracht te verkrijgen, dienen de volgende drie vragen te worden gesteld:

1. Welke interventie moet er uitgevoerd worden,
2. bij welke groep mensen, en
3. om welk resultaat te verkrijgen?

Alvorens op deze vragen dieper in te gaan, worden de componenten van het behandel-effect besproken en wordt beargumenteerd waarom men in effectonderzoek een referentie- ofwel een controlegroep nodig heeft.

De componenten van het effect

Het waargenomen effect (WE) van een bepaalde behandeling wordt in principe niet alleen veroorzaakt door het specifieke effect van deze interventie. Er zijn andere factoren die hierbij een rol spelen.

In de eerste plaats kan het gebeuren dat de klacht van een patiënt vermindert of zelfs verdwijnt, zonder dat er een behandelaar aan te pas is gekomen. Het effect wordt in dat geval beïnvloed door de spontane genezing als eindpunt van het 'Natuurlijk' Beloop (NB) van de aandoening. Stel dat een patiënt met keelpijn bij de huisarts komt, deze wijst de pijn aan een beginnende keelontsteking en schrijft het homeopathische middel Tonsiotreen voor. Wanneer de pijn nu verdwijnt en de patiënt geen tekenen van een keelontsteking meer laat

zien, is niet duidelijk of dit het gevolg is van de interventie, of dat de klacht ook zonder interventie verdwenen zou zijn.

Omdat er in de praktijk na identificatie van een (beginnende) aandoening veelal ingegrepen zal worden, is er over het natuurlijk beloop doorgaans weinig bekend. Iedere interventie heeft naast een specifiek effect ook een aspecifiek effect. Meestal bedoelt men hiermee het verschijnsel dat een patiënt zich beter voelt, louter en alleen doordat van de behandeling een gunstig effect wordt verwacht, ongeacht welke interventie er uitgevoerd wordt; het Placebo-Effect (PE).

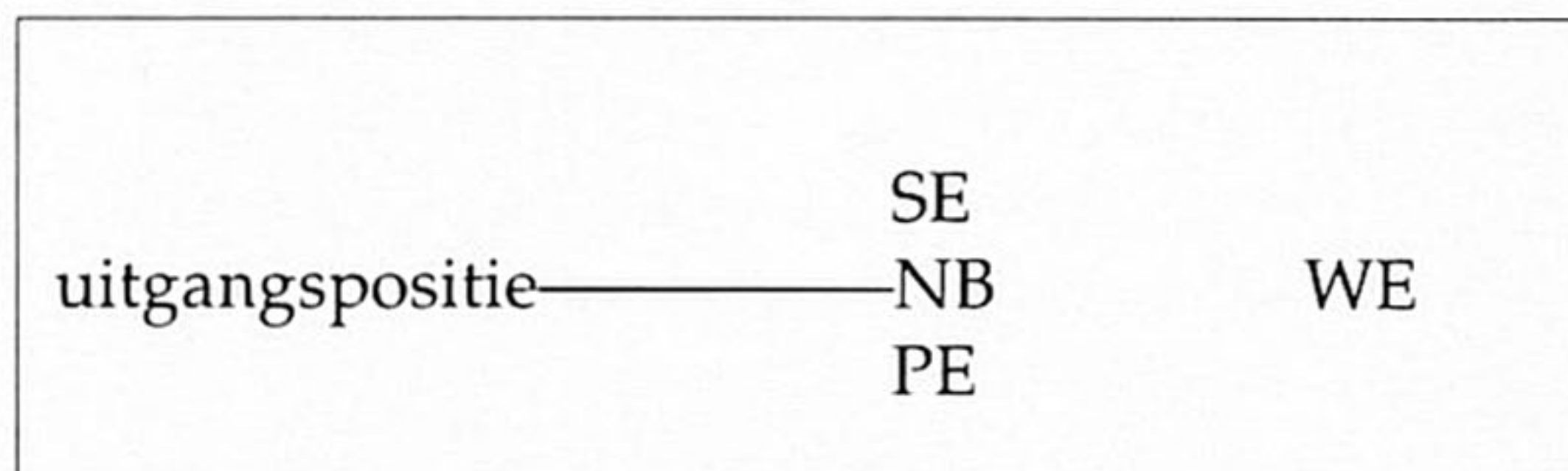
Zo rapporteerde Beescher (4), op grond van 14 onderzoeken naar het placebo-effect van behandelingen voor onder andere post-operatieve wondpijn, pijn bij angina pectoris, zeeziekte en angst, een verbetering in 30 tot 58% van, de met een placebo, behandelde patiënten. Op grond van het feit dat de positieve werking van het placebo-effect reëel is, zou men kunnen concluderen, dat het onderscheiden van een specifiek en een aspecifiek effect slechts van academisch belang is. Om theoretische en om wetenschappelijke redenen is men echter toch veelal in het Specifieke Effect (SE) van de interventie geïnteresseerd. Helaas is dit deel van het Waargenomen Effect (WE) slechts indirect meetbaar. Het verband tussen de genoemde drie factoren en het waargenomen effect kan als volgt worden weergegeven:

$$WE = SE + NB + PE$$

Logischerwijs volgt hieruit dat het specifieke effect van een bepaalde interventie gelijk is aan het waargenomen effect, verminderd met de som van het natuurlijk beloop en het placebo-effect.

$$SE = WE - (NB + PE)$$

Dit kan ook in het eerder gepresenteerde figuur tot uitdrukking worden gebracht.



figuur 2.

Wil men de specifieke werking van een bepaalde interventie vaststellen, dan dient dus eerst de bijdrage van het natuurlijk beloop en het placebo-effect aan het waargenomen effect bepaald te worden.

Door gebruik te maken van een controlegroep, waarbij geen interventie wordt uitgevoerd, kan



men de invloed van het natuurlijk beloop bepalen. Zo kreeg in een gerandomiseerd onderzoek naar het effect van acupunctuur bij nekpijn (5) de onderzoeksgroep een acupunctuur-behandeling volgens de klassieke meridiaan theorie.

In de controlegroep werd de reeds jaren gevolgde, weinig succesrijke therapie voortgezet. Veranderingen die tijdens het onderzoek in deze controlegroep optreden, zou men kunnen interpreteren als het natuurlijk beloop van nekpijn. Door een controlegroep in te stellen, waarin de bestudeerde interventie een placebo is, kan bepaald worden in welke mate deze interventie het placebo-effect overstijgt.

Junnila (6) heeft bijvoorbeeld een onderzoek gedaan naar het effect van acupunctuur bij verschillend gelocaliseerde chronische pijn. Door middel van randomisatie kreeg de onderzoeksgroep een acupunctuurbehandeling op de meridiaanpunten en de controlegroep een pseudo-acupunctuurbehandeling. Deze placebo-behandeling bestond uit het prikken met de nagel in de huid, ongeveer 1 cm van de traditionele acupunctuurpunten, teneinde een pijnlijk gevoel te veroorzaken. Na twintig minuten werd dit herhaald om de deelnemer het gevoel te geven dat de 'naald' verwijderd werd. Vrijwel alle controlepersonen hadden de indruk gekregen dat ze een echte acupunctuurbehandeling gekregen hadden. Deze pseudo-acupunctuurbehandeling werd mede als een adequate placebo-behandeling beschouwd, omdat de huid op deze wijze niet geperforeerd wordt.

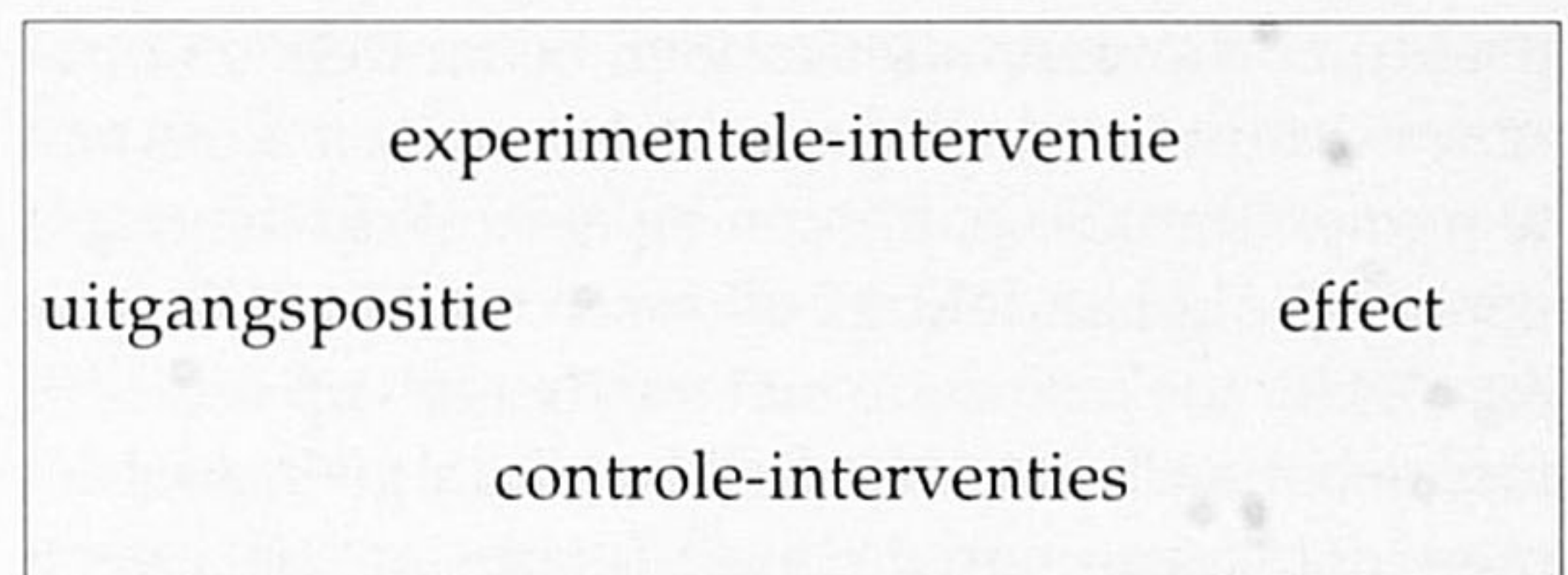
Onderzoeksopzet

Na dit pleidooi voor het gebruik van een controlegroep richten we nu de aandacht op de drie vragen, die de grote lijnen van de opzet van effectonderzoek definiëren. Het antwoord op elk van de vragen is sterk afhankelijk van de onderzoeksvraag die men zich stelt. Daarbij zal duidelijk worden dat men bij het beantwoorden van de betreffende vragen ongelukkige keuzen kan maken, die er toe leiden dat het effect van een interventie door nog weer andere factoren dan het natuurlijk beloop en het placebo-effect vertekend raakt. Het instellen van een controlegroep op zich biedt hiervoor geen oplossing.

Welke interventies zullen met elkaar worden vergeleken?

Er bestaan in principe nogal wat mogelijkheden (controle-interventies), waarmee men de bestudeerde interventie (experimentele-interventie) kan vergelijken.

Wanneer de onderzoeksvraag luidt, of een bepaalde interventie werkt (effectiviteit), is een ver-



figuur 3.

gelijking van die interventie met geen behandeling en/of een placebo-behandeling op zijn plaats.

Zo werd in een onderzoek naar het analgetisch effect van acupunctuur bij gastroscopie (7), een vergelijking gemaakt tussen de klassieke vorm van acupunctuur en een placebo-behandeling, waarin de naalden 1 cm van de klassieke meridiaanpunten geplaatst waren. Stelt men zich daarentegen de vraag, of een bepaalde interventie betere resultaten oplevert in vergelijking tot andere interventies, dan zou men kunnen kiezen voor een vergelijking met een standaard interventie, waarvan de effectiviteit reeds bewezen, dan wel algemeen geaccepteerd is. Wanneer een dergelijke standaard interventie niet bestaat, is er nog de mogelijkheid om de experimentele-interventie te vergelijken met een of meer redelijke alternatieven. In dat geval is het raadzaam tevens een vergelijking met een placebo-interventie in het onderzoek op te nemen, doordat van al deze interventies veelal nog onduidelijk is of ze wel een specifiek effect hebben.

In een onlangs geformuleerd voorstel aangaande een effect-onderzoek naar de effectiviteit van manuele therapie en fysiotherapie bij niet-specifieke rug- en nekklachten (8) wordt er bijvoorbeeld een vergelijking gemaakt tussen manuele therapie, fysiotherapie, behandeling door de huisarts en een placebo-behandeling.

Welke interventies er ook uitgevoerd worden, tijdens de uitvoering zijn zorgvuldigheid en nauwkeurigheid geboden. Dit is van belang om überhaupt een effect te kunnen bewerkstelligen, maar ook om de kans op vertekening van het effect te verkleinen dan wel te voorkomen.

Op de eerste plaats moeten alle interventies vakkundig uitgevoerd worden, in de zin dat ze hun veronderstelde werk naar behoren kunnen doen. Dit betekent, dat ze op het juiste moment, in voldoende sterkte en frequentie, onder de juiste omstandigheden, met de geëigende hulpmiddelen en door een ervaren persoon worden toegepast. In dit kader moet ook de afweging plaatsvinden tussen een nauw omschreven, vastliggende interventie en een interventie die in uitvoering kan variëren, afhankelijk van de reactie van de deelnemer. Het zou natuurlijk niet juist zijn, wanneer de interventie in de onderzoekssituatie volstrekt anders is dan in de



praktijk. Wanneer alternatieve behandelaars opereren vanuit het standpunt dat het onjuist is om aan te nemen dat alle patiënten bij eenzelfde diagnose een gelijke behandeling behoeven, moet men hen in een onderzoekssituatie niet vastleggen op een uniforme behandeling. Op zich hoeft dit geen problemen op te leveren in de interpretatie van het effect, mits er goed geregistreerd wordt wat er tijdens de interventie gebeurt.

Zo kregen in een onderzoek naar de effectiviteit van acupunctuur bij lage rugpijn (9), de acupuncturisten alle vrijheid om hun normale beleid te volgen. Voor alle in een effect onderzoek betrokken verschillende interventies geldt, dat men er voor moet oppassen dat ze wel echt van elkaar verschillen.

Zo riep een onderzoek (10), waarin getracht werd vast te stellen dat een homeopathische verdunning boven de zogenaamde grens van Avogadro (waarin geen molecuul van de oorspronkelijke stof meer aanwezig is), meer dan alleen een placebo-effect veroorzaakt, een aantal zeer kritische reacties op. In dit onderzoek vergeleek men de werking van een dergelijke verdunning; 30cc stuifmeel verdunning van verschillende grassoorten, met een placebo-ogenschijnlijk dezelfde tabletten gedrenkt in een vloeistof van 90% alcohol. Uit de reacties kwam de mening naar voren dat op die manier twee placebo's met elkaar werden vergeleken (11,12) en dat de placebo niet correct gekozen was (13).

Vervolgens dient te worden nagegaan, of co-interventies de invloed van de experimentele en controle-interventies beïnvloeden. Het gaat hier om interventies, die niet direct een onderdeel vormen van de te toetsen interventies, maar wel op hetzelfde moment worden uitgevoerd en die het effect daarvoor mede kunnen beïnvloeden. Deze co-interventies kunnen aan de onderzoeksinterventies gekoppeld zijn en doelbewust gegeven worden, vlak vóór, tijdens, of vlak na de toepassing van de onderzoeksinterventies, om een gezamenlijk effect te bewerkstelligen.

Een voorbeeld hiervan is een post-operatieve kankertherapie, die toegepast wordt ongeacht wat er bij de operatie gevonden wordt. In een gerandomiseerd onderzoek (14) naar de werking van Rheumasan-baden in vergelijking met modderbaden bij pijn in het lendengebied werden bijvoorbeeld per individu tegelijkertijd andere behandelingen voortgezet (fysiotherapie, medicamenten). Hierdoor kon de waargenomen verbetering alleen op de totale behandeling betrokken worden en was het onmogelijk om precies te onderscheiden welk aandeel de badentherapie en welk aandeel de andere therapeutische maatregelen hierin hadden geleverd.

In een onderzoek naar het effect van manuele therapie bij pijnklachten in de nek, arm of hand (15) kre-

gen deelnemers in de onderzoeksgroep, wier nek te pijnlijk was voor een dergelijke behandeling, een injectie met methylprednisolone of een mengsel van lignocaïne en hydrocortisone. Ook kunnen co-interventies uitgevoerd worden als reactie op het effect veroorzaakt door de onderzoeksinterventies. Hieronder valt een post-operatieve kankertherapie, die alleen toegepast wordt, indien er metastasen gevonden worden. Tenslotte kunnen de co-interventies gericht zijn op een ander doel. Om hetzelfde voorbeeld nog verder uit te werken, kan men hier denken aan antibiotica of steroïden, die als ondersteunende therapie gegeven worden aan de patiënt die met chemotherapie behandeld wordt. Er zijn nog twee andere fenomenen die verband houden met de interventies die aandacht verdienen, omdat zij eveneens het effect van de interventies kunnen vertekenen.

Ten eerste dient voor alle deelnemers te worden vastgesteld, in hoeverre zij zich aan de interventie houden. Oorzaken van slechte compliance dienen zoveel mogelijk achterhaald te worden. In het hierboven aangehaalde onderzoek naar de pijnstillende werking van Rheumasanbaden bij lumbaalischialgie werd precies nagegaan hoeveel baden men genomen had, waarbij het maximaal mogelijke aantal als 100%-compliance gedefinieerd werd (14).

Een tweede bron van zorg vormen aanvullende interventies buiten het onderzoek. Daar men er in effectonderzoek naar streeft, groepen met elkaar te vergelijken die slechts verschillen qua ingestelde interventie, dient er op toe gezien te worden dat deelnemers niet tevens een andere interventie ondergaan. Alle soorten vertekening die op kunnen treden tijdens de uitvoering van de te vergelijken interventies, vatten wij hier samen onder de term Observatie Bias (OB). Hiermee wordt de formule als volgt:

$$SE = WE - (NB + PE + OB)$$

Bij wie zullen de interventies worden uitgevoerd?

Het gaat hier om de vraag, wie er, op grond van welke toelatings- en uitsluitingscriteria, in de onderzoekspopulatie opgenomen zal worden. Een nauwkeurige specificatie van deze criteria helpt de onderzoeker bij het maken van een rationele keuze ten behoeve van de vorming van een homogene onderzoekspopulatie. De lezer van het onderzoeksverslag wordt hierdoor in staat gesteld om te bepalen, of door de gekozen samenstelling van de populatie de probleemstelling beantwoord kan worden en in hoeverre het mogelijk is de resultaten van het onderzoek te extrapoleren naar de eigen praktijk.



Uitsluitingscriteria kunnen bijvoorbeeld zijn:

- a) het fysiek niet in staat zijn de interventie te ontvangen,
- b) het niet vatbaar zijn voor een verandering in de aandoening ten gevolge van een interventie,
- c) de onmogelijkheid gevolgd te worden in de tijd dat het onderzoek duurt (voor zover dit vooraf ingeschat kan worden),
- d) het gelijktijdig ontvangen van andere behandelingen die het effect kunnen beïnvloeden,
- e) een extreem goede of extreem slechte prognose hebben wat betreft de aandoening en
- f) het niet in staat kunnen zijn toestemming te verlenen.

Het vormen van groepen die gelijk zijn in hun uitgangspositie, is een voorwaarde om het effect van verschillende interventies eerlijk met elkaar te kunnen vergelijken. Het zou immers een vertekend beeld van het effect geven, wanneer er in een bepaalde groep meer mensen zitten, voor wie de kans op het verwachte effect van de behandeling groter is dan in een andere groep, ongeacht de interventie die men krijgt.

Dit noemt men vertekening door prognostische verschillen (susceptibility bias). Deze verhoogde kans op een goed resultaat kan samenhangen met demografische variabelen, zoals bijvoorbeeld leeftijd en geslacht. Minstens zo belangrijk zijn klinische variabelen zoals bijvoorbeeld de aard, het stadium, de oorzaak en de duur van de aandoening waarvoor de interventie plaatsvindt, andere lichamelijke of psychische aandoeningen, alsmede voorafgaande behandelingen.

Deze variabelen (confounders) vertekenen het effect van de interventie wanneer ze zowel een determinant zijn van het beoogde effect, als ongelijk verdeeld zijn over de verschillende interventies. Doordat ze niet het gevolg zijn van een foutieve onderzoekopzet of van een onjuiste wijze van gegevensverzameling, kunnen ze in een onderzoek veelal niet worden voorkomen. Men kan echter voor de variabelen waarvan men de verstoringe werking kent, in principe corrigeren tijdens de data-analyse. Onbekende confounders vallen onder het natuurlijk beloop van een aandoening en hiervoor kan men slechts met behulp van randomisatie proberen te corrigeren. Tenslotte is de kans op succes hoogst waarschijnlijk groter voor deelnemers die er van overtuigd zijn, dat zij baat zullen vinden bij de interventie. Mocht het zo zijn dat het succes van een behandeling in hoge mate afhangt van het geloof in die behandeling, zoals dit in de alternatieve geneeswijzen veelvuldig gesteld wordt, dan zou een effectonderzoek bij voorkeur moeten zijn gericht op uitsluitend deze subgroep. Hier zou men dan immers het grootste effect mogen verwachten.

Dit betekent vervolgens dat, wanneer men in een dergelijke onderzoekspopulatie geen effect vindt, dit zeer waarschijnlijk ook geldt voor elke andere populatie. Het feit dat de onderzoeksgroepen verschillen qua prognose, kan ook veroorzaakt worden door beslissingen die bewust of onbewust door zowel behandelaars als patiënten genomen worden. Elke behandelaar zal over het algemeen die interventie uitvoeren, waarvan hij verwacht dat deze de meeste voordelen oplevert.

Een gevolg hiervan in een effect-onderzoek zou kunnen zijn, dat de patiënten die in de ene groep terecht komen, verschillen van de patiënten in de andere groep(en) qua prognose en dus in hun kans op het beoogde effect. Deze vorm van vertekening, 'confounding by indication', kan zich voordoen in alle gevallen waarin het mogelijk is dat de ingeschatte prognose de kans beïnvloedt om aan een bepaalde interventie te worden toegewezen. Ook de patiënt kan een bepaalde reden, die wel of niet te maken kan hebben met de klacht (kosten, locatie, reputatie, laatste kans, etc.), op een bepaald moment (eerste symptomen, terminale fase, etc.), besluiten om de hulp in te roepen van een bepaalde behandelaar (regulier, alternatief). Er is dan dus sprake van een zelf-selectie door de patiënt. De verschillende redenen waarom mensen met een bepaalde klacht zich aanmelden als vrijwilligers voor een onderzoek, kunnen er dan ook toe bijdragen dat er een vertekening van het effect ontstaat, wanneer die redenen de kans op het behalen van het effect vergroten of verkleinen.

In een gerandomiseerd onderzoek naar het effect van manuele therapie bij acute lage rugpijn is getracht deze bias ten gevolge van de selectie door de patiënt te minimaliseren door vooraf het onderzoek niet in de publiciteit te brengen (16). Wat de omvang van de onderzoekspopulatie betreft, wordt hier volstaan met te constateren dat deze groot genoeg moet zijn om eventuele relevant geachte verschillen tussen het effect van de interventies te ontdekken. Ook hier biedt een homogene groep deelnemers voordeel, doordat dan minder mensen deel hoeven te nemen.

Selectie Bias (SB) is een overkoepelende term die vaak gebruikt wordt, voor alle fouten die gemaakt kunnen worden tijdens het specificeren en het selecteren van de onderzoekspopulatie.

Dit leidt tot een uitbreiding van de formule:

$$SE = WE - (NB + PE + OB + SB)$$

Welke effectmaten zijn relevant?

Met het uitvoeren van een interventie wordt een bepaald effect beoogd. Wil men dit effect kunnen meten, dan is het noodzaak precies te omschrijven



welk effect het hoofddoel van de interventie is en vervolgens te definiëren in welke variabelen het effect kan worden uitgedrukt en welke hiervan de belangrijkste zijn. Uit de literatuur blijkt bijvoorbeeld dat er een reeks van onderzoeken gedaan zijn naar de werking van ginseng bij ouderen (17). Eén van de beoogde effecten van ginseng is een verhoging van de belastbaarheid en weerstand van ouderen. Dit effect is op verschillende manieren geoperationaliseerd zoals in een vermindering van allerlei subjectieve klachten (moeheid, slapeloosheid, nervositeit, etc.), verbeteringen in cognitieve functies en in de relatie met de omgeving, een groter reactie-vermogen op licht- en geluidprikkels, een toegenomen concentratie, een beter herinneringsvermogen en meer ondernemingszin.

De uitkomst die men het belangrijkste achtte, was over het algemeen de invloed van ginseng op het voorkomen van allerlei subjectieve klachten. Hieruit mag al blijken, dat de onderzoeker niet per definitie die effectparameters behoeft te kiezen die het meest objectief en met grote precisie kunnen worden vastgesteld.

Een volgende stap is het kiezen van de meetinstrumenten voor deze geoperationaliseerde variabelen en de bepaling van de tijdstippen waarop de effectmeting dient plaats te vinden. In de ginsengstudies werd onder andere gebruik gemaakt van standaardvragenlijsten, psychologische tests en psychomotorische tests. Om het effect van ginseng vast te kunnen stellen, werden de scores van de deelnemers op de variabelen in de uitgangspositie vergeleken met hun scores vanaf een aantal weken nadat zij ginseng waren gaan gebruiken. Er werd dus een vergelijking gemaakt tussen de situatie voor de interventie en de van elkaar te onderscheiden momenten, waarop een meetbaar effect zich kan hebben voorgedaan, het effect nog toegenomen kan zijn en de tijd waarover het effect na de behandeling kan blijven voortduren.

Meerdere metingen gedurende een follow-up zijn met name van belang, wanneer de aandoening, waarvoor men de interventie uitvoert, chronisch van aard is, onder andere om te bepalen of een gunstig effect eerder optreedt in vergelijking met een andere behandeling. In een onderzoek naar het effect van het plaatsen van naalden in neurofysiologisch gekozen motore-punten, werd een laatste follow-up meting verricht gemiddeld 30 weken na de beëindiging van de behandeling om het lange termijn-effect te kunnen bepalen (18).

Een derde stap is de keuze van de analysetechnieken en van de maat waarin men de uitkomsten uitdrukt. Kort gezegd, wordt dit in belangrijke mate bepaald door onder andere de gehanteerde meetinstrumenten, de wijze waarop de onderzoekspopulatie is gekozen (wel/niet a-select) en

het karakter van de relatie tussen de interventie en het effect. Voor een uitvoerige uitleg van het statistisch toetsen op significantie en de voorwaarden en aannames die daarbij een rol spelen, wordt verwezen naar de epidemiologische handboeken, die zijn vermeld in de literatuurlijst.

Tenslotte dienen ook andere dan de beoogde effecten, waaronder bijwerkingen, vermeld te worden. Verschillen tussen het waargenomen effect en het specifieke effect kunnen ook veroorzaakt zijn door fouten bij het meten van de effectparameters. Dit soort meetfouten (TM) kan een toevallig karakter hebben en daardoor het resultaat van de metingen op een onvoorspelbare wijze beïnvloeden. Aan de andere kant kunnen er ook systematische fouten bij het meten worden gemaakt, waardoor allerlei vormen van bias ontstaan.

Eerder zijn reeds observatie bias en selectie bias genoemd als overkoepelende termen. Een verwante vorm van vertekening, en meestal gezien als onderdeel van observatie bias, kan zich voordoen in alle situaties, waarbij de mogelijkheid om een bepaald effect vast te stellen verschillend is voor de te vergelijken onderzoeksgroepen. Immers, om een bepaald effect te kunnen ontdekken dienen er een drietal activiteiten te worden ondernomen;

1. evalueren van de deelnemers,
2. bepaling van het effect en
3. interpretatie van het effect.

Om hierbij vertekening te voorkomen behoren alle deelnemers even regelmatig gecontroleerd te worden en daarbij dezelfde mate van aandacht te krijgen. Iedereen dient met dezelfde meetinstrumenten en gelijke nauwkeurigheid onderzocht te worden en de beoordeling van het effect mag niet subjectief gekleurd worden. Dit laatste kan bijvoorbeeld gemakkelijk gebeuren, wanneer de observator van het effect in het oordeel wordt beïnvloed door de wetenschap welke interventie een deelnemer heeft gekregen. En verwante vorm van vertekening, en eveneens vallend in het domein van selectie-bias, kan ontstaan wanneer de groep, die uiteindelijk in de analyse wordt meegenomen, afwijkt van de oorspronkelijke onderzoekspopulatie.

Door allerlei omstandigheden kunnen deelnemers uit de onderzoekspopulatie verdwijnen gedurende het onderzoek. Gelukkig vormt dit niet altijd een bron van vertekening, bijvoorbeeld wanneer de factoren die tot de uitval hebben geleid, geen verband houden met het effect van de interventie. Aan de andere kant zou het effect van een interventie flink overschat kunnen worden, wanneer alle personen die zich tussentijds hebben teruggetrokken uit ontevredenheid met de behandeling, niet betrokken worden in de analyse. Het is dan ook wenselijk, voor zover mogelijk, redenen van non-re-



spons en 'loss-to-follow-up' te achterhalen. Als regel zou men bij effectonderzoek niet af moeten wijken van het 'intention-to-treat' principe; elke deelnemer dient, wat er ook gebeurt, in de analyse te worden meegenomen, als behorende bij de interventie welke bij het begin van het onderzoek is toegewezen. Uiteindelijk levert dit de volgende definitieve formule op:

$$SE = WE - (NB + PE + OB + SB + TM)$$

- SE = Specifiek Effect
- WE = Waargenomen Effect
- NB = Natuurlijk Beloop inclusief onbekende confounders

•PE = Placebo Effect, het niet-specifieke effect

•OB = Observatie Bias

Vertekening tijdens de uitvoering van de interventies

Vertekening door de werking van co-interventies

Vertekening door de verschillen in kans of detectie

•SB = Selectie Bias

Vertekening door prognostische verschillen, veroorzaakt door:

a) confounding by indication (door de behandelaar, door de patiënt)

b) onderkende confounders

Vertekening door wijziging in het originele cohort (loss-to-follow-up, non-respons)

TM = Toevalige Meetfouten (ruis)

CONTROLE OP VERTEKENING

We hebben al vastgesteld dat men voor het natuurlijk beloop en het placebo-effect kan corrigeren door gebruik te maken van controlegroepen. De volgende vraag is dan, of men de verschillende vormen van vertekening, de overige componenten in de formule, kan voorkomen, dan wel hoe men hiervoor kan corrigeren.

Experimenteel onderzoek

Het kenmerkende van experimenteel onderzoek is een, op een van te voren vastgesteld plan gebaseerde toewijzing van de deelnemers aan de experimentele groep of aan de controlegroep.

Uitgaande van de veronderstelling dat de onderzoeker interventiegroepen met elkaar wil vergelijken die slechts verschillen in interventie, maar tegelijkertijd een gelijke verdeling hebben, wat betreft hun uitgangspositie, de kans op ontdekking van het effect, en de wijze waarop de interventie is uitgevoerd, staan de onderzoeker twee krachtige instrumenten ter beschikking om dit doel te bereiken. Door randomisatie en blinding wordt de best mogelijke controle bereikt over alle bekende en onbekende variabelen die het effect beïnvloeden.

Randomisatie

Uit het voorgaande is gebleken dat het specifieke effect door allerlei factoren kan worden vertekend. Het is daarom van cruciaal belang dat de onderzoeksgroepen vergelijkbaar zijn met betrekking tot hun prognose, los van een interventie-effect. Het hulpmiddel bij uitstek om dit te bereiken is randomisatie. Dan bepaalt alleen het toeval aan welke interventie een bepaalde persoon wordt toegewezen. In een onderzoek naar het effect van een vitamine-C supplement op het herstel van patiënten met decubitus via een verhoogde collageen-vorming, vond randomisatie plaats op het geboortjaar van elke deelnemer (19). Aldus hoopt men tevens de onbekende en slecht meetbare confounders gelijk te verdelen over de groepen. Uit deze voorzichtige formulering blijkt overigens dat, hoewel randomisatie een essentieel bestanddeel is van een experiment, dit geen garantie biedt voor een daadwerkelijke gelijke verdeling van confounders. Vooral wanneer het onderzoek vrij kleine groepen betreft, kunnen er door het toeval toch ongelijke verdelingen ontstaan. Extra bescherming van de geldigheid van het onderzoek kan de onderzoeker inbouwen, door de totale onderzoekspopulatie eerst te verdelen in subgroepen ten aanzien van bepaalde kenmerken die het effect sterk kunnen beïnvloeden (stratificeren) en personen daarbinnen vervolgens a-select toe te wijzen aan de experimentele dan wel aan de controle-interventie.

Deze subgroepen worden dan vooraf geformeerd op basis van veronderstelde prognostische factoren, bijvoorbeeld leeftijd, geslacht en ziektestadium. Een dergelijke stratificatie vooraf (groep-matching) in de praktijk is slechts voor een beperkt aantal variabelen mogelijk.

Zo werden in een onderzoek naar het effect van ginseng op onder andere het algemeen welzijn, het reactievermogen en de longfunctie, de deelnemers voorafgaande aan de randomisatie gestratificeerd op leeftijd en geslacht (20). Hoewel uit statistisch oogpunt minder efficiënt, kan de onderzoeker ook achteraf, in de analyse, op eventueel ongelijk verdeelde, bekende confounders stratificeren. Een voorwaarde daarbij is, dat deze confounders in het onderzoek gemeten zijn. In de denkwereld van veel alternatieve geneeswijzen beschouwt men de mens als een uniek wezen waarvoor in principe nooit een identieke controle-persoon te vinden is. Dit is echter juist een extra argument om te randomiseren, gezien het grote aantal onbekende confounders dat kennelijk een rol speelt. Daarnaast roept het proces van randomisatie dikwijls vragen op ten aanzien van de ethische aanvaardbaarheid ervan.

In de alternatieve kringen wordt regelmatig het standpunt verdedigd dat randomisatie niet mogelijk is, vanwege het feit dat een patiënt bewust voor



een alternatieve behandeling kiest. Hier tegenover zou men kunnen stellen dat het onethisch is om patiënten de keuze te laten tussen therapieën waarvan het effect nog onvoldoende duidelijk is. Wanneer aan een aantal voorwaarden wordt voldaan, zijn er naar de mening van de auteurs geen bezwaren tegen randomisatie. Ten eerste moet er sprake zijn van 'informed consent'. Elke deelnemer behoort te zijn ingelicht met betrekking tot de verschillende interventies die in het onderzoek met elkaar vergeleken zullen worden, alsmede over het feit dat er gerandomiseerd zal worden. Op basis hiervan geeft elke deelnemer al dan niet toestemming.

Ten tweede dient de totale onderzoekspopulatie samengesteld te zijn uit personen die, welke interventie er ook bij hen uitgevoerd wordt, hierdoor op grond van de huidige inzichten, geen gevaar lopen. Een derde voorwaarde is, dat nog onvoldoende bekend moet zijn welke van de interventies het meest effectief is.

Een laatste voorwaarde is natuurlijk dat alle deelnemers het recht hebben om zich terug te trekken uit het onderzoek. Wanneer zij echter hebben deelgenomen aan het randomisatieproces, betekent dit dat zij definitief zijn opgenomen in de onderzoekspopulatie. Dit is de consequentie van het volgen van het eerder genoemde 'intention-to-treat' principe.

Blindering

Het doel waarnaar men streeft met blinderen, is te maskeren welke interventie door welke deelnemer wordt ondergaan. Idealiter zou hiermee bereikt moeten worden dat zowel de deelnemer, de behandelaar, de observator van het effect en degene die de data analyseert, niet weten welke interventie bij wie is uitgevoerd. Dit zal echter in de praktijk voor de ene partij makkelijker te realiseren zijn dan voor de andere partij en bovendien zal afhankelijk zijn van het type interventie.

Een voorbeeld; in een onderzoek naar het effect van acupunctuur bij osteo-musculaire pijn kreeg de experimentele groep een acupunctuur behandeling waarbij de klacht-specifieke meridiaanpunten werden aangeprikt. De controlegroep kreeg een niet-klacht-specifieke acupunctuurbehandeling (21). Blindering van de patiënt om vertekening als gevolg van het selectieve vertrouwen van een patiënt in een bepaalde interventie te voorkomen, is in dit geval hoogst waarschijnlijk geen probleem. Wanneer er gebruik gemaakt kan worden van een placebo-interventie, is blindering van de patiënt cruciaal, omdat de wetenschap dat men een 'nep'-behandeling krijgt, het placebo-effect tot nul kan reduceren. In een gecontroleerd onderzoek naar het effect van manuele therapie bij chronische nekpijn,

werden de deelnemers aan het onderzoek verdoofd door middel van een injectie met diazepam om blinding van de patiënt te bewerkstelligen (22). Om na te gaan of de blindering succesvol was geweest, vroeg men de deelnemers achteraf, of ze wisten welke behandeling ze hadden gekregen en of ze zich konden herinneren manuele therapie ondergaan te hebben. Voordat men ertoe kan besluiten een placebo op te nemen in een effectonderzoek, zal men zich eerst af dienen te vragen of het gebruik van een placebo-interventie wel gerechtvaardigd is, en zo ja, of een dergelijke interventie praktisch uitvoerbaar is.

De eerste vraag concentreert zich rond een aantal ethische problemen. Het valt te verdedigen dat een placebo aanvaardbaar is, wanneer de deelnemers in de placebo-groep gemiddeld naar verwachting niet veel slechter af zullen zijn dan bij een al geaccepteerde aanpak. Een sine qua non is in alle gevallen, het vooraf informeren van elke potentiële deelnemer die toestemming dient te geven om aan het onderzoek mee te doen, omtrent de kans op een placebo-behandeling.

Ter verduidelijking van de tweede vraag is het belangrijk zich goed te realiseren, dat een placebo-effect alleen dan goed te meten is, wanneer men niet weet dat men een 'nep'-behandeling krijgt. Hieruit volgt de eis dat de placebo-interventie ogenschijnlijk niet verschillend is van de experimentele, 'echte' interventie.

Dit houdt in, dat beide interventies even overtuigend dienen te worden uitgevoerd en in soort, aantal en duur van handelingen op elkaar behoren te lijken. Wanneer de experimentele interventie in de vorm van een pil of injectie gegeven wordt, is een placebo gemakkelijker te realiseren dan wanneer het om bepaalde handgrepen gaat, zoals die onder andere plaatsvinden in de manuele therapie. Tegelijkertijd kan men zich afvragen, of er daadwerkelijk geen specifieke werking uitgaat van een placebo-behandeling.

Zo zou in het acupunctuur-onderzoek waar, in de controlegroep, de placebo-behandeling bestond uit het plaatsen van naalden op punten die men als volkomen onjuist beschouwde, deze behandeling wellicht een specifieke werking kunnen hebben vergelijkbaar aan het exact op de meridiaanpunten prikken (23).

Blindering van de behandelaar om vertekening tijdens de uitvoering van de interventies (OB) te voorkomen, is vaak moeilijk te realiseren, doordat deze weet welke behandeling er gegeven wordt. In hun onderzoek hebben Weintraub et al. (21) hiervoor de volgende oplossing bedacht. Een acupuncturist bepaalde voor alle patiënten, zowel in de experimentele, als de controlegroep, waar geprikt moest worden.



Vervolgens plaatste een arts, niet op de hoogte met acupunctuur-theorieën en met de klacht van de patiënt, de naalden volgens de aanwijzing van de acupuncturist. Blindering van de observator, om te voorkomen dat de kans op het beoogde effect verschillend is voor de experimentele en de controlegroep (OB), kan worden gerealiseerd door bijvoorbeeld een andere persoon dan de behandelaar het effect te laten bepalen.

In een gecontroleerd onderzoek naar het effect van manuele therapie bij chronische nekpijn (22) werd de effectmeting verricht door een arts die niet betrokken was geweest bij de behandeling en tijdens de meting niet wist, welke patiënt manuele therapie had gekregen. Soms is de patiënt zelf observator van het effect en in dat geval, is blindering van de patiënt ook om die reden gewenst.

In een onderzoek naar het effect van acupunctuur in vergelijking tot een placebo-behandeling bij migraine (24), was de effectmeting volledig gebaseerd op een hoofdpijn-dagboek bijgehouden door de deelnemers. Hierdoor werden subjectieve gegevens verzameld omtrent de ernst, frequentie, gebruik van pijnstillende middelen, alsmede over de duur van de hoofdpijnaanval. Blindering van de deelnemers is in een dergelijk geval een vereiste om het specifieke effect van deze middelen te kunnen bepalen.

Blindering van degene die de data analyseert, om vertekening tijdens de interpretatie van de resultaten te voorkomen, kan volledig bereikt worden door de toegepaste interventies zo te coderen dat ze niet herkenbaar zijn. Men dient wat betreft het blinderen, zo goed mogelijk na te gaan in hoeverre men hierin geslaagd is, om daar vervolgens bij de presentatie en interpretatie van het onderzoek rekening mee te houden.

TER AFSLUITING

In dit artikel is getracht te beargumenteren, waarom men, bij onderzoek naar een verondersteld verband tussen een interventie en een effect, uit dient te gaan van een gecontroleerde onderzoeksopzet en waarom een therapeutisch experiment waarbij gebruik wordt gemaakt van randomisatie en blindering, het sterkste bewijs levert van het bestaan van een dergelijk verband. In de wetenschap is het experiment dan ook de 'gouden standaard'. In de opzet van elk effectonderzoek, en het maakt geen verschil of dit nu het reguliere of het alternatieve circuit betreft, zou men ernaar moeten streven zoveel mogelijk aan de strenge methodologische regels van het experiment te voldoen. Een alternatieve methodologie en statistiek bestaat niet en heeft ook niet ontwikkeld te worden. Onderzoekingen naar de waarde van alternatieve ge-

neeswijzen, waarvan er in dit artikel een aantal gepresenteerd zijn, tonen aan dat de huidige onderzoeksmethoden die in de geneeskunde worden gehanteerd, toepasbaar en geschikt zijn.

We hopen hiermee een bijdrage te leveren aan de beëindiging van een nutteloze en onnodige discussie over de manier waarop de alternatieve geneeswijzen onderzocht dienen te worden in de daadwerkelijke uitvoering van effectonderzoek op dit terrein. Voor de lezer die zich verder wil verdiepen in de experimentele onderzoeksopzet, zijn hieronder in de algemene literatuurlijst een aantal suggesties gegeven.

LITERATUUR

1. Mulder J.H. Alternatief en regulier. Protocol voor een onderzoek naar de waarde van een omschreven alternatieve therapie in de behandeling van patiënten met een bepaalde aandoening. NTIG 15, jg 3 (1986) 294-298.
2. Aakster C.W. Commentaar op het protocol voor onderzoek naar de waarde van alternatieve therapieën van J.H. Mulder. NTIG 15, jg 3 (1986) 299-308.
3. Aakster C.W. (Begeleidingscommissie Onderzoek Moerman-Methode) 4e-Protocol voor een onderzoek naar de waarde van de Moerman-Methode in de behandeling van kankerpatiënten. 1983.
4. Beecher H.K. Measurement of subjective responses: Quantitative effects of drugs. New York: Oxford, 1959.5. Coan R.M., Wong G.,
5. Coan P.L. The acupuncture treatment of neck pain: a randomized controlled trial. *Amer.J.Chin.Med.* 9 (1982) 326-332.
6. Junnila S.Y.T. Acupuncture therapy for chronic pain. *Amer.J.Acupuncture* 10 (1982) 259-262.
7. Cahn A.M., Carayon P., Hill C., et al. Acupuncture in gastroscopy. *Lancet* 1 (1978) 182-183.
8. Bouter L.M., Knipschild P.G. Het experiment als paradigma voor interventie-onderzoek in de fysiotherapie. *Ned.T.Fysiotherapie* 97 (1987) 56-61.
9. Coan R.M., Wong G., Ku S.L., et al. The acupuncture treatment of low back-pain: a randomized controlled study. *Am.J.Chin.Med.* 8 (1980) 2, 181-189.
10. Reilly D.T., Taylor M.A., McSharry C., Aitchison T. Is homeopathy placebo-response? *Lancet* 2 (1986) 881-885.
11. O'Keeffe D. Is homeopathy a placebo-response? (Letter to the editor) *Lancet* 2 (1986) 1106-1107.
12. Khan M.F. Is homeopathy a placebo-response? (Letter to the editor) *Lancet* 2 (1986) 1107.
13. Skrabanek P. Is homeopathy a placebo-response? (Letter to the editor) *Lancet* 2 (1986) 1107.
14. Kuban G., Weber K.-M., Horbach L., Arnold A., Hansen J. Klinische Prüfung der analgetischen Wirksamkeit von Rheumabädern bei Lumboschialgien. *Therapie Woche* 35 (1985) 2026-2040.
15. Howe D.H., Newcombe R.G., Wade M.T. Manipulation of the cervical spine-a pilot study. *J.Royal Coll.Gen.Practit.* 33 (1983) 574-579.
16. Godfrey C.M., Morgan P.P., Schatzker J. A randomized trial of manipulation for low-back pain in a medical setting. *Spine* 9 (1984) 301-304.
17. Knipschild P.G. Ginseng: pep of nep? (1986) Ter publicatie aangeboden.
18. Gunn C.C., Chir B., Milbrandt W.E. et al. Dry needling of muscle motor points for chronic low-back pain: a randomized clinical trial with long term follow-up. *Spine* 5 (1980) 279-291.



19. Taylor T.V., Rimmer S., Day B., Butcher J., Dymack I.W. Ascorbic acid supplementation in the treatment of pressuresores. *Lancet* 2 (1974) 544-546.
20. Forgo I., Kayasseh L., Staub J.J. Einfluss eines standardisierten Ginseng-Extraktes auf das Allgemeinbefinden, die Reaktionsfähigkeit, Lungenfunktion und die gonadalen Hormone. *Die Medizinische Welt* 32 (1981) 751-756.
21. Weintraub M., Petursson S., et al. Acupuncture in Musculoskeletal pain: methodology and results in a double-blind controlled clinical trial. *Clin. Pharmac. and Therap.* 17 (1975) 248.
22. Sloop P.R., Smith D.S., Goldenberg E., Dore C. Manipulation for chronic neck-pain: a double blind controlled study. *Spine* 7 (1982) 532-535.
23. Godfrey C.M., Morgan P. A controlled trial of the theory of acupuncture in musculoskeletal pain. *J. Rheumatol.* 5 (1978) 121-124.
24. Dowson D.I., Lewith G.T., Machin D. The effects of acupuncture versus placebo in the treatment of headache. *Pain* 21 (1985) 35-42.

ALGEMENE LITERATUUR

- Kleinbaum D.G., Kupper L.L., Morgenstern H. *Epidemiologic research: principles and quantitative methods*. Belmont (CA): Lifetime Learning Publications, 1982.
- Fletcher R.H., Fletcher S.W., Wagner E.H. *Clinical Epidemiology-the essentials*. Baltimore: Williams & Wilkins, 1982.
- Shapiro S.H., Louis T.A. (Eds.). *Clinical trials: Issues and approaches*. Statistics: Textbooks and Monographs, Vol.46, New York: Marcel Dekker, Inc., 1983.
- Feinstein A.R. *Clinical epidemiology: the Architecture of Clinical Research*. Philadelphia: W.B. Saunders Company, 1985.
- Sackett D.L., Haynes R.B., Tugwell P. *Clinical epidemiology: a basic science for clinical medicine*. Boston: Little, Brown and Company, 1985.

Adres vakgroep Epidemiologie

*Drs. L.M. Bouter
p/a Rijksuniversiteit Limburg
Postus 616
6200 MD MAASTRICHT*